DOCUMENT RESUME

ED 209 652 CS 006 364

AUTHOR Verhoeven, Ludo T.
TITLE Diagnostic Evaluation in the Initial Tuition of Reading and Spelling.
INSTITUTION Centraal Inst. voor Toetsontwikkeling, Arnhem (Netherlands).
PUB DATE Aug 81
NOTE 13p.; Paper presented at the Annual Meeting of the European Conference on Reading (2nd, Joensuu, Finland, August 2-5, 1981).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Beginning Reading; Formative Evaluation; Learning Processes; Primary Education; *Reading Diagnosis; *Reading Instruction; *Reading Tests; Spelling Instruction; *Test Construction; *Test Validity

ABSTRACT

Reporting on the construction and validation of instruments for the evaluation of students in the first stages of reading and spelling instruction, this paper first discusses how teachers themselves can determine a student's learning progress with the help of evaluation instruments. Distinguishing between formative evaluation and diagnostic evaluation, the paper then provides a discussion of the construction of a set of tests that can be used in both ways. The paper argues that the tests make it possible for teachers to apply psychological principles of learning to read to the design of interventions that will improve student reading performance. (FL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED209652

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Cito Centraal Instituut voor Toetsontwikkeling

Diagnostic evaluation in the initial tuition of reading and spelling

- Paper for the Second European Conference on Reading -
2 - 5 August 1981, Joensuu, Finland

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Ludo Th. Verhoeven
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

July 1981
Ludo Th. Verhoeven

CS206956

Full adress of the author;
c/o National Institute for Educational Measurement
P.O. Box 1034
6801 MG Arnhem
Netherlands

1 Introduction

This paper is a report of the construction and validation of instruments for the evaluation of pupils in the first stages of reading and spelling instruction. It first discusses how a teacher himself can determine a pupil's progress in the learning process with the help of evaluation instruments. A distinction is made here between formative evaluation on the one hand and diagnostic evaluation on the other hand. Then a brief survey is given of the construction of a set of tests wich can be used both formatively and diagnostically. With this survey it will be illustrated how with information processing task analysis psychological theory of learning to read can be linked to instructional practice.

2 Evaluation

Evaluation is generally understood te mean deciding on the value of something. In education and more in paricular in the field of pupils' behaviour evaluation has come to mean assessing the effects of learning with a view to make decisions . Depending on the type of decisions one wishes to make, we may distinguish five objectives of evaluation:

- formative evaluation: the process with the help of which the progress of instruction can be periodically adjusted;
- summative evaluation: the process with the help of which an assessment can be given after the conclusion of a certain phase in instruction;
- diagnostic evaluation: the process with the help of which specific difficulties in the learning process of weak pupils can be traced;
- placement evaluation: the process with the help of which it is decided which type of training within a given type of school is best suited for a pupil;
- selective evaluation; the process with the help of which it is decided which type of school is best suited for a pupil.

This paper is concerned with the construction and validation of instruments which aim at both formative

and diagnostic evaluation. These types of evaluation are concerned with the determination of progress within the instuctional situation. In such a situation it is the teacher's task to teach a group of pupils a certain amount of subjectmatter which has a more or less logical structure. In designing an adequate teaching system a teacher should, to begin with, determine what goals of instruction he wants to reach in a given period. Next he has to plan the instructional device that will make it possible to reach those goals. Finally he has to decide what evaluation procedures he wants to use.

Using evaluation procedures is important for the teacher. They enable him to adapt the instructional proces to the specific needs of the individual pupils. The teacher will have to check regularly whether the objectives formulated beforehand, have been attained by the pupils. With the help of observation, oral turns and selfmade tests the teacher can get a global impression of the progress of his pupils. However, such evaluation procedures do not give exact information on where the instruction has to be adjusted. This can be done with standardized tests, and more in particular with domain-referenced tests. A domain-referenced test is not interpreted by comparing the effects of learning of pupils with those of an external group of reference, but by assessing those results in terms of the domain of subject matter the testcontent refers to. Performances on these tests give the teacher more precise information on where the instruction has to be adjusted. With this information the teacher can adapt the teaching process by giving additional instruction. This is called formative evaluation.

If, however, a pupil continues to have difficulties in learning, a more extensive diagnosis is required. The teacher will then have to find out the nature and the possible causes of these difficulties. This implies: finding out to what degree the processes basic to a task are disfunctioning. This type of evaluation is called diagnostic evaluation. It implies interpreting the effects of learning in terms of the processes basic to them An adequate diagnosis can be made with the help of diagnostic tests. Such tests show the teacher how the teaching process can best be adjusted to the specific needs of weak pupils.

3 Diagnostic-tests in the first stages of reading and spelling instruction

Cito has developed a new set of tests for the evaluation of pupils in the first stages of reading and spelling instruction. These tests are meant to be used both formatively and diagnostically.

The tests aim at measuring three skills: reading words, writing words and reading texts. A total of six tests has been developed: two tests for word-decoding, two tests for word-dictation and two tests for reading comprehension. These tests are set in the first year of the reading and spelling instruction. The precise moment to set these tests is not fixed: they are set the moment the subject matter that is incorporated in the test has been dealt with. Thus the degree to which a pupil has mastered the subject matter is examined. Furthermore each test allows for the results of the pupil to be interpreted diagnostically. In this way teachers are informed about the underlying factors that account for weak performances at tests.

Constructing the tests the following procedure has been adopted: As a first step some four readingcurricula have been analysed (of these four, two are being used by 90 per cent of the teachers). For each of these curricula it is examined which objectives are dealt with successively. Then it is determined which objectives mark the end of a certain phase in the process of learning to read in each of these curricula. An example of such an objective is reading all the CVC words. Such objectives each determine a certain domain of subject matter. On the basis of such a domain a finite domain of test items can be defined. By making a sample from such a domain of items one arrives at a domain referenced test. Such a test can be seen as a formative evaluation instrument that can be set the moment a teacher has dealt with the subject matter from the domain concerned. The test then indicates to what degree pupils have mastered the subject matter dealt with by the teacher. For a sufficient command of the subject matter pupils should score at least 85 per cent correctly.

As a second step in the construction process procedures have been examined which could diagnose weak pupils, i.e. the non-masters on the tests. On the basis of theories on learning to read we have tried to link up the subject

matter incorporated in the tests with psychological processes underlying the tasks in the tests. With the help of information processing task analysis (Resnick, Beck, 1975) it was analysed which subskills play a role in reading and writing words and short texts. Also an analysis was made of the errors in the various tests. Then for each test it was investigated what the connection was between the occurence of certain types of errors in the test and the mastery (or absence of mastery) of subskills that underly the test performance. It then proved that in a number of cases the occurence of certain types of errors could be related to an insufficient command of specific subskills which play an important role in the initial stages of learning to read and write. Thus the tests can be seen as diagnostic evaluation instruments. By analysing the pupils' errors in the tests the teacher will be informed about the degree to which his pupils have mastered the underlying subskills.
By means of two of the tests this construction process will now be illustrated.

4 Construction of a diagnostic test for word-reading

Word-decoding test 1 is a test that is meant to measure pupils' ability to read words of the CVC type. In the reading curricula the instruction with the use of this type of words is finished after four months.
In order to measure this ability with the help of a test the following method of operationalisation has been chosen. The test consists of 48 items, each consisting of a more or less complex drawing and three words going with it. The pupil has to read the words and to indicate which of these words goes with the object in the drawing. This operationalisation is in accordance with natural reading behaviour in which meaning must be given to a written code without oral expression.
For a sufficient command a pupil must score 85 per cent correctly. If a pupil scores less, it is advised to analyse the errors made by the pupil. How this can be done will now be discussed.
In the items only one of the words is correct. This word is of the CVC type. To the correct word two incorrect words have been added systematically: one word in which one

letter from the correct word has been replaced by another letter, and one word which differs from the correct word in one of the following aspects:
a letter has been added or left out or the letter order has been changed (see figure 1).

☐ pair
☐ hair
☐ air

☐ pear
☐ bear
☐ pearl

Figure 1: two samples of Word decoding 1 (adjusted to English)

Thus a distinction is made between two possible types of error. This choice of types of error has been made after a thorough analysis of the task required from the pupils when making this test: (see appendix 1)
A pupil reads the first word and he can arrive at an interpretation in two ways: on the basis of direct recognition or on the basis of recoding (La Berge, Samuels, 1976). Because 'sight-words' have been avoided in the tests, the pupil will be left with recoding. The recoding process presupposes three steps: the word is usually analysed in graphemes; then seperate graphemes (groups of graphemes) are coupled with phonemes (groups of phonemes) and finally these phonemes are combined to make up the spoken word. Weak readers appear to have problems with the last two steps (Shankweiler Liberman, 1972). The subskills of grapheme - phoneme coupling and phonemic synthesis have not been mastered sufficiently. Pupils who are bad at coupling graphemes with phonemes often interchange one letter with another letter; pupils who are bad at phonemic synthesis often omit letters, add letters or reverse words.
In the word decoding test it can be analysed to what degree weak readers have problems with the grapheme - phoneme coupling, with phonemic synthesis or with both subskills.
Once the test had been constructed (it consists of two parallel versions) it was analysed extensively. The reliability of both versions proved to be very high (KR-20>.90); average Rit>.50. In order to control the content validity of the test a fit-measure of the latent-trait-model of Rasch was determined. Model-

fit was found for both versions of the test. Two procedures were used to find out about construct validity. First a number of correlations were measured: test performance and the teacher's judgement, test performance and performance on a reading aloud-test; and test performance and existing tests that measure the subskills grapheme-phoneme coupling and phonemic synthesis.
All the correlations proved to be greater than .50. Also the internal consistency of patterns of errors was examined. For this the KR-20 formula was used (one score for type of error 1 and one score for type of error 2). These coefficients proved to be greater than .80 for both versions of the test. Finally, it was checked whether pupils who made many errors of one type were also weak at the supposed underlying skill. This was indeed the case.

5 Construction of a diagnostic test for word writing

Word dictation 1 is a test meant to measure the spelling ability of pupils at words of the CVC, CCVC and CVCC type. In the reading curricula the correct writing of these types of words has been paid much attention to after 6 months. The test is based on 32 words which have a phonetic spelling. In other words: these words have a great grapheme-phoneme consistency. These words are offered orally. Also, the context of the words is given with the help of a picture in the test-book (figure 2).

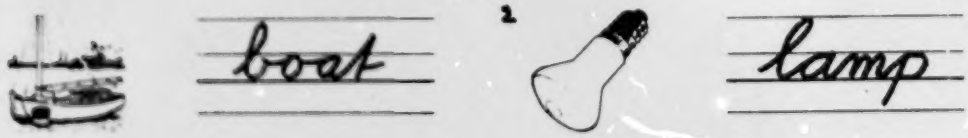


Figure 2: two samples of Word dictation 1 (adjusted to English)

For a sufficient command a pupil must score 80 percent correctly. If a pupil has written less than 80 percent of the words correctly, it is advised to analyze the errors made by the pupils. How this can be done is shown by the following procedure.
Each error is given one of the following codes:

Code	Error type	Possible cases
1	insertion of a letter	
2	omission of a letter	
3	reversal of words	
4	reversal of diphthongs	oe-eo,ui-iu,eu-ue,ie-ei.
5	interchange of vowels	a-aa,a-o,aa-oo,eu-u,etc.
6	mirroring of letters	b-d,b-p,t-f,k-k,etc.
7	interchange of sound-similar consonants	v-f,s-z.
8	interchange of other consonants	b-k,f-g,k-l,etc.
9	mutilation of a word	two/more errors in one word

In order to gain a better insight into the underlying factors which determine the various types of errors, an analysis of the pupils task when writing a word must be made. Information-processing task-analysis shows the following: (see appendix 2)

The pupil hears the orally presented word and he can arrive a written code in two ways: on the basis of direct recognition or on the basis of encoding (Frith,1979). As 'sight-words' have again been avoided, the pupil is left with encoding. The encoding process presupposes two steps. First the word is broken down into separate phonemes (groups of phonemes); then these are coupled with the correct graphemes (groups of graphemes). More or less at the same time the word is being written down. Weak spellers seem to have mastered these subskills (phonemic analysis and the phoneme-grapheme coupling) insufficiently. Pupils who are bad at phonemic analysis are known to easily omit or insert letter or to reverse words (in casu error codes 1-3) Pupils who are bad at phoneme-grapheme coupling often interchange one letter with an other letter (in casu error code 4-8) In other words task analysing shows that the error types distinguished before refer back to an insufficient command of these two subskills. With the help of the word-dictation testit can be analysed to what degree pupils have difficulties with phonemic analysis, with the phoneme-grapheme coupling or with both subskills.

Once the test had been constructed (it consists of two parallel versions) it was analysed extensively. The reliability of both versions proved to be very high (KR-20> .90; average R_{it} > .50) Fit of the latent-trait-model of Rasch was found with both test versions,

so that it can be concluded that the items are a homogeneous group. As far as the construct validity is concerned a number of correlations were measured: test performance and the teacher's judgement and test performance and tests that measure the subskills phonemic analysis and phoneme-grapheme coupling. All the correlations proved to be greater than .55. Also the internal consistancy of the presupposed clusters of errors was examined. For both versions the KR-20 proved to be between .60 and .80. Finally it was shown that pupils who made many errors of one type were also weak at the supposed underlying skill.

6 Conclusion

With the help of two examples I have tried to illustrate how tests may contribute to the diagnosis of learning difficulties in the field of learning to read.
With these tests it will be possible to apply psychological principles of learning to read to the design of interventions that will facilitate the acquisition of desirable levels of reading performance. The tests indicate which pupils have difficulties in learning to read. They also give an indication in the nature of these problems and in the way these problems can be solved.
It is of great importance that the teacher can use the tests himself. In this way the teacher can himself improve the quality of instruction, especially for those pupils who are in danger of falling behind in the teaching process.

7 References

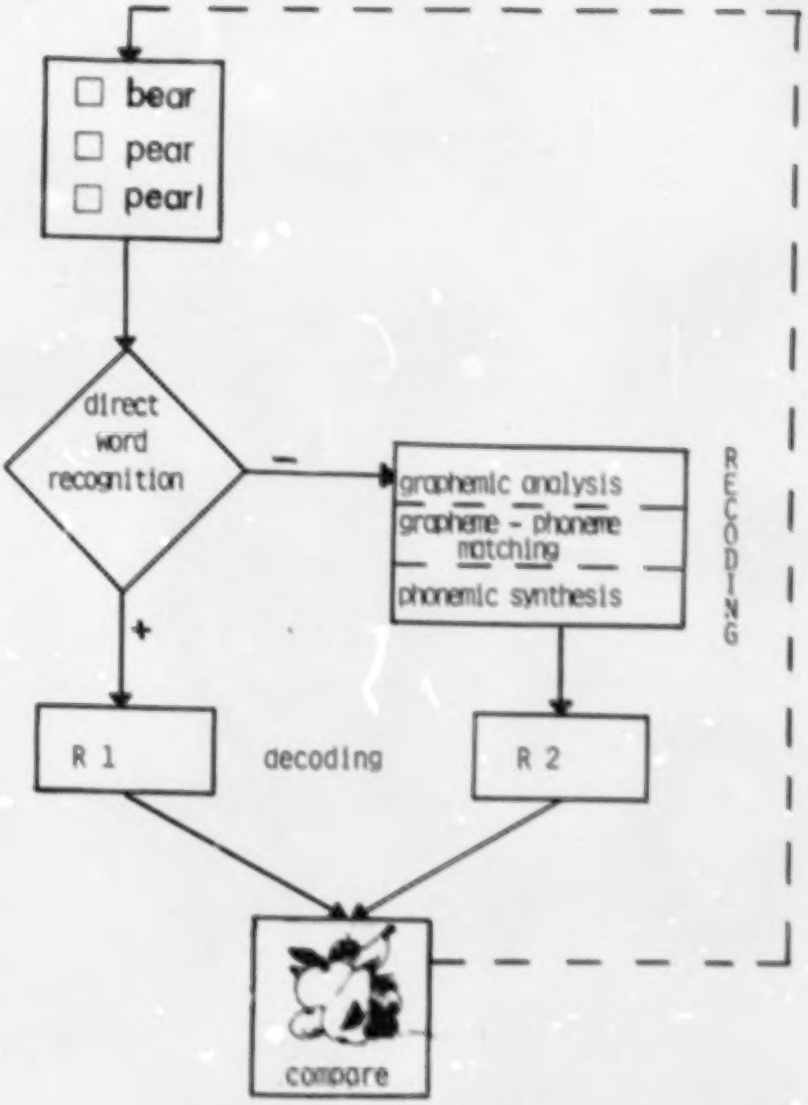
Frith, U., Cognitive processes in spelling, Academic press, London, 1979.

La Berge, D. and S. Samuels Toward a theory of automatic information processing in reading.
In: Singer, S. and R. Ruddell (eds.) Theoretical models and processes of reading. Newark Delaware, IRA, 1976.

Resnick, L. and I. Beck Designing instruction in reading: integration of theory and practice.
In: Guthrie, J. (ed.) Aspects of reading acquisition. Baltimore, John Hopkins University Press, 1975.

Shankweiler, D. and I. Liberman, Misreading, a search for causes. In: Kavanagh, J. and I. Mattingly (eds.) Language by ear and by eye. Cambridge, MIT Press, 1972.

Task analysis Word-decoding 1



Appendix 2

Task analysis Word-dictation 1